

Non-association between Neo-5 personality tests and multiple mini-interview

Kulamakan Kulasegaram · Harold I. Reiter · Willi Wiesner ·
Richard D. Hackett · Geoffrey R. Norman

Received: 5 June 2009 / Accepted: 20 November 2009 / Published online: 15 December 2009
© Springer Science+Business Media B.V. 2009

Abstract Most medical schools attempt to select applicants on the basis of cognitive and non-cognitive skills. Typically, interpersonal skills are assessed by interview, though relatively few applicants make it to interview. Thus, an efficient paper and pencil test of non-cognitive skills is needed. One possibility is personality tests. Tests of the five factor model of personality, and in particular the factor of conscientiousness, has proven effective in predicting future job performance. Can it serve as a screen for admissions interviews? In particular, correlation with the multiple mini-interviews (MMI) is of interest since the latter is a well validated test of non-cognitive skills. A total of 152 applicants to Michael G. DeGroote School of Medicine at McMaster completed the Neo-5 personality test voluntarily in advance of their admissions interviews. Correlations were calculated between personality factors and grade point average (GPA), medical college admissions test (MCAT) and MMI. No statistically significant correlation was found between personality factors and cognitive (GPA, MCAT) measures. More surprisingly, no statistically significant correlation was found between personality factors, including conscientiousness, and the MMI. Personality testing is not a useful screening test for the MMI.

Keywords Multiple mini-interview · Personality test · Neo-5 test · Screening test · Medical school · Admissions · Validity

Introduction

Most medical schools attempt to select students for both non-cognitive and cognitive abilities. While cognitive measurement is relatively straightforward using measures such as the medical college admissions test (MCAT) scores and undergraduate grade point average (GPA), schools have been hard pressed to find valid and reliable measures of interpersonal and communication skills (Kreiter and Kreiter 2007; White et al. 2009). The traditional

K. Kulasegaram · H. I. Reiter (✉) · W. Wiesner · R. D. Hackett · G. R. Norman
McMaster University, Hamilton, ON, Canada
e-mail: reiterh@mcmaster.ca

answer has been to use personal essays and interviews. Unfortunately, these have demonstrated poor reliability and validity (Kreiter et al. 2004; Eva and Reiter 2004; Albanese et al. 2003; Edwards et al. 1990).

One promising measure is the multiple mini-interview (MMI), which has been shown to have acceptable reliability and validity (Eva and Reiter 2004). The MMI has been adopted by over a dozen Canadian and American medical schools, in large part because of its ability to predict pre-clerkship OSCE performance, clerkship performance, and national licensing examination scores in the non-biological sciences (ethics, communication, population; Eva et al. 2004a; Reiter et al. 2007). In the MMI candidates are evaluated through a number of interviewing stations that test diverse topics and skills such as ethics, collaboration, and knowledge of the medical system (Eva et al. 2004b). The median reliability of the MMI from multiple studies has been reported as 0.73 and it is correlated 0.43 with the number of stations passed in OSCE-based licensing examination (Eva et al. 2009). Despite this, the medical school applicant pool size precludes MMI usage for more than a few 100 candidates, simply on the basis of feasibility. Thus a screening tool is required at the initial stages of application.

One possibility is a self-completed scale directed at identifying candidates with desirable personality characteristics. Personality testing has a long history in the human resources literature (Barrick and Mount 1991; Barrick et al. 1993). Some dental schools have moved towards using personality testing to determine whether applicants are suitable and some medical schools have piloted the use of a medicine—specific personality battery although there is no evidence as yet of its validity (Lumsden et al. 2005; Smithers et al. 2004; Bezner and Boucher 2001). Personality tests can be administered without the intensive resources required of the MMI. As such, if a personality test can be shown to be related to non-cognitive abilities, it could possibly function as a screening tool for the MMI since they are both intended to measure the same, or similar, domains.

Early personality testing studies relied on the Myers-Briggs type indicator (MBTI) or the California Psychological Inventory as the preferred tests. Despite its long history, the MBTI has been found lacking because of its poor validity and reliability (Boyle 1996; Pittenger 2005; Zemke 1992). Further, the MBTI and the CPI are not based on the most current advances in personality theory. The five factor model of personality (or colloquially the big five) is currently the most popular taxonomy for personality characteristics and has been used in a variety of settings including medical admissions (Barrick and Mount 1991; Barrick et al. 1993; Goldberg 1990; Chamberlain et al. 2005). The model aims to assess the most important personality factors and their associated facets: neuroticism, extroversion, openness to experience, agreeableness, and conscientiousness. Neuroticism is the tendency to experience poor emotional adjustment and negative emotions such as anxiety, insecurity and hostility. Extroversion reflects sociability, assertiveness, positive emotions, zeal and sensation seeking; openness to experience is the disposition to be imaginative, nonconforming, unconventional and autonomous; Agreeableness is the tendency to be trusting, compliant, caring and gentle; and conscientiousness is comprised of achievement and dependability (Judge et al. 2002). The five factors each offer six sub-facets, for a total of thirty facets of personality (Goldberg 1990).

Testing of the five factor model has generally shown value in work settings in predicting various features of job performance (Barrick and Mount 1991). Generally, the factor that is most often correlated with performance is conscientiousness (Barrick and Mount 1991; Barrick et al. 1993). Explorations of the five factor model in medical and dental school admissions confirm this relationship. Lievens et al. found a significant difference between successful and unsuccessful medical students in three facets of conscientiousness:

dutifulness, achievement striving, and self-discipline, although the overall multiple correlations of the five scales were not strong ($R = 0.25$ for first year, 0.17 for second year, 0.22 for third year; Lievens et al. 2002). Ferguson et al. used five factor testing in comparison to grades, reference letters, and personal statements and found that of the five factors, conscientiousness was the strongest predictor of preclinical performance. The correlation ranged from 0.38 to 0.49 with preclinical grades, exceeding the predictive ability of previous grades for preclinical performance (ranging from 0.10 to $.30$). In contrast to its ability to predict preclinical grades, the predictive validity of conscientiousness with clinical grades was a good deal lower, ranging from 0.08 to 0.21 (Ferguson et al. 2003). Chamberlain et al. used five factor testing at a Canadian dental school and found that conscientiousness predicted Year 1 and Year 3 scores with correlations at 0.23 and 0.47 , $p < .05$ (Chamberlain et al. 2005). This study also found sub-facets of conscientiousness to correlate with grades and performance. Similar studies have been conducted in medical schools comparing personality testing with peer assessment of tutorial performance (Luh et al. 2007). There is not perfect consistency across studies however; Smithers et al. (2004) found that openness to experience was the best predictor of success in dental school, though the correlation was negative. Extroversion was also predictive, but conscientiousness was not.

Despite the somewhat encouraging results of these studies, use of the Neo-5 or any personality measure appears under-conceptualized on two grounds. First, the initial rationalization for the use of personality measures is to tap into the non-cognitive domain of interpersonal and communication skills. The use of grades as a criterion variable, and the consequent emergence of conscientiousness as the most important trait, does not provide any evidence of the utility of Neo-5 to predict non-cognitive dimensions, where the real need lies. Second, even if appropriate non-cognitive outcomes were identified (for example, communication with patients, leadership, moral and ethical behaviour), it is not at all clear which aspects of personality or combination of traits might be most predictive and hence most useful to score for admissions. No theory maps personality traits onto optimal clinical performance. In the absence of theory, we must rely on the summation of many correlations of personality variables with many outcomes as indicating a possible linkage between personality testing and valid measurement of non-cognitive performance.

For this reason, an examination of the relationship between Neo-5 and MMI does begin to explicitly examine the possible association between personality and non-cognitive performance since, as we have indicated, there is ample evidence that the MMI is a valid predictor of non-cognitive outcomes in medical school. If it can be shown that one or more dimensions of the Neo-5 are highly correlated with the MMI, the Neo-5 could possibly be used as a screening test, since compared to the significant resources demanded by the MMI, most common five factor personality tests only require completion of an easily accessible questionnaire format. Previous research does suggest that conscientiousness and agreeableness would predict professional behaviour in dental students (Chamberlain et al. 2005). An association between the MMI and the Neo-5 would not only provide a plausible screening tool for the MMI, but also begin to test a theory relating specific personality dimensions to non-cognitive abilities.

The study described in this paper is an attempt to examine the correlation between MMI and five-factor personality test scores from the Neo-5 in medical school applicants. We hoped that an examination of the pattern of correlations between personality traits and MMI scores would begin to define a theory of the relationship between personality and non-cognitive ability. In order to examine independence from traditional cognitive measures, correlations between personality factors and GPA were tested. The MMI and Neo-5

were also tested against scores on the MCAT verbal reasoning as it is another predictor of clerkship performance.

Research questions

1. Can a standardized measure of the “big five” personality dimensions serve as a screening test of non-cognitive skills, as assessed by the MMI?
2. What components of the big five are related to non-cognitive skills as assessed by the MMI?

Methods

Sample and procedures

In the 2008 medical school admissions cycle at McMaster, 538 people were assessed by the MMI interview. The mean GPA of this group was 3.84 (SD 0.12), where possible scores range from 0 to 4. Of this group, 357 had written the MCAT, see <http://www.aamc.org/students/mcat/>) at some point during their admissions process. The MCAT is a comprehensive test consisting of individual scales measuring knowledge of physical sciences, biological sciences, verbal reasoning, and writing skill. Scores range from 1 to 15 for each sub-scale. All interviewees were invited to complete the Neo-5 personality test. Of the 538, 152 completed the questionnaires (28.3%). Individual sub-scales had slightly differing levels of completion with neuroticism, $n = 151$, extroversion $n = 146$, openness to experience $n = 150$, agreeableness $n = 149$, and conscientiousness $n = 152$. The study was approved by the Hamilton Health Sciences Research Ethics Board.

Measures

MMI

MMI methodology has already been extensively described (Eva et al. 2009). In the 2008 medical school admissions run at McMaster, 12 stations were used, with sub-domains in communication, collaboration, critical thinking, ethics, personal statement, and understanding of the health care system. Each station ran for 8 min of total interview time and was assessed by a single interviewer with scoring based on a scale of 1–10 for an overall score. Scores were summed across stations to generate the total score. No subscores were generated since any subscore, derived from relatively few stations, would be unreliable.

Neo-5

The Neo-5 personality test was originally designed by McCrae and Costa (1994) and underwent subsequent revision to longer and shorter forms (McCrae and John 1992). The short version of the test was used which consists of 60 items which ask the user to express varying levels of agreement with each item (McCrae and Costa 1994). The Neo-5 has been validated and tested for reliability in several populations including the college age population the United States (McCrae and Costa 1994). Item scores are collected as raw

summations and were then converted into T-scores for each of the five factors. Scores from all Neo-5 personality tests were totalled by a single rater with 20 tests randomly selected to be re-scored to ensure that there was agreement on scoring. Data collected from applicants were then entered along with the MMI data.

Analysis

Pearson's correlations were used to examine the relation between Neo-5 subscales and MMI scores. We also examined relation to GPA to confirm absence of a relationship to cognitive measures and with MCAT Verbal Reasoning as this has been shown to be a good predictor of communication ability and of clerkship performance (Kulatanga-Moruzi and Norman 2001). The criterion for significance was 0.05, however this was adjusted for the number of correlations ($n = 21$) using Bonferroni correction ($\alpha = 0.002$).

Results

Table 1 shows mean and standard deviation of GPA, MMI, MCAT verbal reasoning and Neo-5 scores for participants.

Table 1 Descriptive statistics

	Mean	Std. deviation	<i>N</i>
GPA	3.84	0.12	538
MCAT verbal reasoning	9.60	1.62	357
MMI	63.94	12.98	538
Neo-5			
Neuroticism	16.04	7.76	152
Extroversion	34.38	5.52	146
Openness to experience	32.88	6.61	150
Agreeableness	37.92	5.01	149
Conscientiousness	39.16	5.67	152

Table 2 Comparisons of means between non-volunteers and volunteers

	Non-volunteers	Volunteers
GPA		
<i>N</i>	387	151
Mean	3.86	3.81
SD	0.117	0.135
MCAT verbal reasoning		
<i>N</i>	273	84
Mean	9.64	9.60
SD	1.57	1.62
MMI total		
<i>N</i>	151	387
Mean of total score	63.48	65.11
SD	13.65	11.05

Table 3 Correlations between personality subtests

	Neuroticism	Extroversion	Openness to experience	Agreeableness	Conscientiousness
Neuroticism					
Extroversion	−.25				
Openness to experience	.05	.30			
Agreeableness	−.19	.30	.25		
Conscientiousness	−.40	.13	−.01	.17	

Table 4 Correlations between personality subtests and outcome measures

	GPA	MCAT verbal	MMI total
GPA		.07	.06
MCAT verbal	.07		.10
MMI	.06	.10	
Neuroticism	.12	.13	.02
Extroversion	−.23	−.01	.10
Openness to experience	−.20	.20	.22
Agreeableness	.01	−.13	.12
Conscientiousness	.02	−.20	.01

Table 2 shows the GPA, MMI, and MCAT verbal reasoning scores for those who did not complete the personality questionnaires, and those who did, to determine if the volunteer sample was systematically biased. An ANOVA of these measures (GPA, MMI, and MCAT) showed a significant difference in GPA scores in favour of the non-volunteers, [$F = 15.68$ ($p < .0001$)], however the difference was only 0.05, which is not sufficient to bias computed correlations. There was no significant difference in MCAT or MMI total scores between the volunteer and non-volunteer groups.

Table 3 shows the correlations among the subscores of the Neo-5. Correlations ranged from $-.40$ to $+.30$ (Mean = .035) with the largest negative correlation between neuroticism and conscientiousness, and the largest positive correlations between extraversion and openness to experience and agreeableness.

Table 4 shows the correlations between personality score and the other admissions measures. No personality variable correlated significantly with MMI total score. Mean correlation was 0.09, and the largest (non-significant) correlation was 0.20, with openness to experience. MMI also did not correlate with GPA or MCAT verbal reasoning.

Discussion

The results show no significant correlation between MMI and Neo-5 on any relevant personality factors. As anticipated, the MMI did not correlate with GPA. Personality scales and performance measures such as MMI represent two very different approaches to assessing the non-cognitive domain; one assessing stable traits, the other behaviours. Nevertheless, the hope of this study was that personality measures would have a role in the

admissions process as a potential predictor of non-cognitive skills. The evidence was not supportive, however.

While ideally, validity studies should examine future performance in medical school and beyond, there are at least three reasons why a concurrent approach, such as used here, may provide unique insight. First, to the extent that a measure is used in selection, this will inevitably result in restriction of range when examining outcomes in medical school, so that a low correlation can be viewed as an inevitable consequence of the process, and can be used to deflect claims of failure. Second, the passage of time from selection to outcome can and does impact on potential association; so again, a low correlation might be seen as inevitable, and not as evidence of poor validity. Neither limitation applies to the concurrent validation of two measures in a sample from the whole applicant pool. Third, since the goal of personality testing, at least as we envision using it, is as a screen for non-cognitive abilities, an examination of its relation to MMI scores is a direct test of this construct.

In that regard, the results from the study are not promising. While there is the inevitable problem of which measure can be viewed as a criterion, the MMI has been previously shown to measure non-cognitive ability and predict performance using a variety of medical school and national licensure outcomes. The exact specifics of this domain are still under investigation but increasingly it seems the MMI measures some concatenation of interpersonal skills, ethical reasoning, and situational judgement (Reiter et al. 2007). Thus, the lack of correlation with the personality test is a concern.

The study has some limitations. It was conducted in a single year at a single medical school. Further, the participation rate was relatively low. While there was a small difference in GPA, we have no reason to suppose that volunteers differed systematically from non-volunteers in such a way that the magnitude of the correlations would be suppressed in a volunteer cohort.

The negative findings of the current study are not inconsistent with the literature. The magnitude of correlations is actually comparable; however published literature may be capitalizing on Type I errors. Most published reports comparing personality factors and performance invariably run numerous correlations, sometimes ranging over one hundred. Naturally, it is expected that every 100 correlations will yield approximately five that are found significant by chance alone. This was the rationale behind the use of Bonferroni correction in the present study.

A second concern is that the general practice in these studies is to provide post hoc explanations for each of the statistically significant correlations found. While corrections for this practice have been made in the human resources literature (Barrick et al. 2001), the same cannot be said for studies looking at dental and medical school applicants. For example, Smithers et al. found little correlation for conscientiousness, selecting openness to experience as the stronger factor for dental school admissions on the basis of correlations with academic performance measures (Smithers et al. 2004). Chamberlain et al. (2005) concluded later that conscientiousness in fact was the stronger factor for prediction. Positive correlations with agreeableness and extroversion and the negative correlations with neuroticism have in turn been used to argue that one factor or combination of factors would prove fruitful in medical selection processes. To lend support to the correlations, one would expect that consistent correlation should be found in independent studies. Comparisons between studies, however, fail to demonstrate consistency, indicating rather a haphazard allocation of those correlations (Smithers et al. 2004; Ferguson et al. 2003; Chamberlain et al. 2005).

There are some exception to this trend. For example, Poole et al. found that conscientiousness correlated from 0.24 to 0.47 with seven criterion measures of academic and

clinical dental school performance (Poole et al. 2007). Of the other personality factors however, only one out of the 28 (less than 5%) was significant. However, as we indicated earlier, if personality does predict academic performance, while perhaps of theoretical interest, it is of little practical value since traditional measures such as GPA and MCAT have stronger relation to outcomes and are available at little cost.

The variability of the correlations from study to study reflects an underlying under-conceptualization of the area. While there is no clear and widely accepted definition of “non-cognitive” skills, there are certainly defensible measures of communication skills, ethical reasoning, etc. But the association between non-cognitive skills and personality traits remains elusive. It may be tempting to presume that extroverted individuals have better interpersonal skills in clinical settings, but this is hardly a defensible theory. The state of the art would be advanced by a clear statement of exactly which personality variables might be expected to correlate with which outcomes, so the inevitable statistical “shot gunning” is avoided.

It also appears that personality measures lack construct validity over time; the predictive validity of conscientiousness and other factors generally decreases or becomes negative with passage of time. For example, Lievens et al. (2002) found six facets of conscientiousness correlated strongly with cognitive domain preclinical Year 1 performance, but the strength and the significance of these correlations became progressively weaker from then on. If personality is viewed as stable and trait-like, then the finding of decaying relation of conscientiousness to subsequent outcomes can be viewed as evidence against the predictive validity of the measure. At best, its value becomes diluted once the student moves beyond the level of novice. Furthermore, personality is also affected by the context in which a personality trait can be expressed, thus potentially losing any inherent value when success is judged based on a wide array of situations and contexts.

Conclusions

The results of the current study have shown essentially no relationship between five personality variables as measured by the Neo-5 and MMI Scores. As a result, the hope of using this particular form personality testing as a screening test of non-cognitive skills was not borne out.

Acknowledgments Funding was provided by SSHRC (Social Sciences and Humanities Research Council) grant (SSHRC Standard Research Grant 410-2004-0541).

References

- Albanese, M., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine*, 78, 313–321.
- Barrick, M. R., & Mount, M. K. (1991). The Big 5 personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9–30.
- Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology*, 8, 715–722.
- Bezner, J. R., & Boucher, B. K. (2001). The influence of personality type on decision making in the physical therapy admission process. *Journal of Allied Health*, 30, 83–91.

- Boyle, G. J. (1996). Myers-Briggs type indicator (MBTI): Some psychometric limitations. *Australian Psychologist*, 30(1), 71–74.
- Chamberlain, T. C., Catano, V. M., & Cunningham, D. P. (2005). Personality as a predictor of professional behaviour in dental school: Comparisons with dental practitioners. *Journal of Dental Education*, 69, 1222–1236.
- Edwards, J. C., Johnson, E. K., & Molider, J. B. (1990). The interview in the admission process. *Academic Medicine*, 3, 167–176.
- Eva, K. W., & Reiter, H. I. (2004). Where judgement fails: Pitfalls in the selection process for medical personnel. *Advances in Health Sciences Education*, 9, 161–174.
- Eva, K. W., Reiter, H. I., Rosenfeld, J., & Norman, G. R. (2004a). The ability of the multiple mini-interview to predict preclerkship performance in medical school. *Academic Medicine*, 10(Suppl), S40–S42.
- Eva, K. W., Reiter, H. I., Trinh, K., Wasi, P., Rosenfeld, J., & Norman, G. R. (2009). Predictive validity of the multiple-mini interview for selecting medical trainees. *Medical Education*, 4, 767–775.
- Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004b). Admissions OSCE: The multiple mini-interview. *Medical Education*, 38, 314–326.
- Ferguson, E., James, D., O’Hehir, F., & Sanders, A. (2003). Pilot study of the roles of personality references, and personal statements in relation to performance over the 5 years of a medical degrees. *British Medical Journal*, 326, 428–431.
- Goldberg, L. R. (1990). An Alternative description of personality the big 5 factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87, 765–780.
- Kreiter, C. D., & Kreiter, Y. (2007). A validity generalization perspective on the ability of undergraduate GPA and the medical college admission test to predict important outcomes. *Teaching and Learning in Medicine*, 19, 95–100.
- Kreiter, C. D., Yin, P., Solow, C., & Brennan, R. L. (2004). Investigating the reliability of the medical school admissions interview. *Advances in Health Sciences Education*, 9, 147–159.
- Kulatanga-Moruzi, C., & Norman, G. R. (2001). Validity of admissions measures in predicting performance outcomes: The contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine*, 14, 34–42.
- Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students’ personality characteristics and academic performance: a five-factor model perspective. *Medical Education*, 36, 1050–1056.
- Luh, S. P., Yu, M. N., Lin, Y. R., Chou, M. J., Chou, M. C., & Chen, J. Y. (2007). A study on the personal traits and knowledge base of Taiwanese medical students following problem-based learning instructions. *Annals Academy of Medicine Singapore*, 36, 743–750.
- Lumsden, M. A., Bore, M., Millar, K., Jack, R., & Powis, D. (2005). Assessment of personal qualities in relation to admission to medical school. *Medical Education*, 39, 258–265.
- McCrae, R. M., & Costa, P. T. (1994). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, 36, 587–596.
- McCrae, R. M., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175–215.
- Pittenger, D. J. (2005). Cautionary comments regarding the Myers-Briggs type indicator. *Consulting Psychology Journal: Practice and Research*, 57, 210–221.
- Poole, A., Catano, V. M., & Cunningham, D. P. (2007). Predicting performance in Canadian dental schools: The new CDA structured interview, a new personality test, and the DAT. *Journal of Dental Education*, 71, 664–676.
- Reiter, H. I., Eva, K. W., Rosenfeld, J., & Norman, G. R. (2007). Multiple mini-interviews predict clerkship and licensing examination performance. *Medical Education*, 41, 378–384.
- Smithers, S., Catano, V. M., & Cunningham, D. P. (2004). What predicts performance in Canadian dental schools? *Journal of Dental Education*, 68, 598–613.
- White, C. B., Dey, E. L., & Fantone, J. C. (2009). Analysis of factors that predict clinical performance in medical school. *Advances in Health Sciences Education*, 14, 455–464.
- Zemke, R. (1992). Second thoughts about MBTI. *Training*, 28, 43–45.